Fourth Aviation Psychology Symposium: II

Multidimensional Scaling Analysis of Simulated Air Combat Maneuvering Performance Data

Donald J. Polzella, Ph.D., and Gary B. Reid, M.A.

Polzella DJ, Reid GB. *Multidimensional scaling analysis of simulated air combat maneuvering performance data.* Aviat. Space Environ. Med. 1989; 60:141–4.

This paper describes the decomposition of air combat maneuvering by means of multidimensional scaling (MDS). MDS analyses were applied to performance data obtained from expert and novice pilots during simulated air-to-air combat. The results of these analyses revealed that the performance of expert pilots is characterized by advantageous maneuverability and intelligent energy management. It is argued that MDS, unlike simpler metrics, permits the investigator to achieve greater insights into the underlying structure associated with performance of a complex task.

MULTIDIMENSIONAL SCALING (MDS) is a family of statistical techniques that attempts to discover the "hidden structure" in data consisting of measures of relatedness among sets of objects (e.g., things, events, stimuli) (11). MDS uses a matrix of proximities among the objects as input and produces an N-dimensional configuration or map of the objects as output. The configuration is derived in such a way that the distances between the objects in the configuration match the original proximity measures as closely as possible. More importantly, the locations of particular clusters of objects are said to reflect whatever dimensions might underlie the proximity measures.

Various proximity measures are suitable for analysis by MDS (15). They include: a) direct judgements of similarity or dissimilarity of objects; b) sortings or clusterings of objects; c) stimulus confusion data; d) interobject latencies and amplitudes; e) communication, social interaction, or volume flow between objects; f) comparisons of rating profiles; and g) intercorrelation matrices. The list of suitable measures suggests that MDS is an extremely flexible procedure. It is not surprising, therefore, that MDS techniques have been applied to a wide variety of research problems in the applied behavioral sciences. Beatty (1), for example, used MDS to map the functional proximity of spontaneous electroencephalographic data obtained from different regions of the cerebral cortex. Medlin and Thompson (9) used MDS to determine the dimensions used by expert Army evaluators when they assessed unit performance in field exercises. Schvaneveldt, *et al.* (12) used MDS to compare experts' and novices' memory structures for critical flight information. Laxar, Moeller, and Rogers (8) used MDS to determine the cognitive organization of submarine sonar information. Some unusual applications of MDS include an analysis of thermal pain (3), cross-media artistic styles (4), the effect of viewing position on the perceived layout of space (14), and the organization of visual memory in the rhesus monkey (10).

This paper describes an application of multidimensional scaling to pilot performance data obtained during simulated air-to-air combat. Performance measurement of air-to-air combat presents formidable problems for the research scientist. An adequate performance metric must be capable of describing the complex interrelationship between position advantage and energy management, which characterizes the air-to-air combat environment. According to Breidenbach, *et al.* (2), existing measurement methods and metrics either do not meet this criterion or have not been adequately tested.

The Database

In the present study, we have applied MDS to data collected by Kelly, *et al.* (5) on the Simulator for Air-to-Air Combat (SAAC), which is located at Luke AFB, AZ. The SAAC consists of two F-4 cockpits (one of which can be operated with flight characteristics similar to those of a MIG-21 aircraft) controlled by a common computer system (Xerox SIGMA 5). Two camera-model systems provide each pilot with a display of the opposing aircraft, which is superimposed on a 296° (horizontal) by 150° (vertical) field of view. Additional motion cues are provided by G-seat and G-suit systems.

Kelly, *et al.* (5) collected their data during 405 one-versus-one (1v1) air combat maneuvering (ACM) free engagements on the SAAC. The subjects were 30 F-4 qualified pilots of varying levels of experience. The data consisted of continuous samples of aircraft system variables, engagement outcomes and events, and evalua-

---

From the University of Dayton, Dayton, Ohio (Dr. Polzella); and Harry G. Armstrong Aerospace Medical Research Laboratory, Wright-Patterson AFB, Ohio (Mr. Reid).

Address reprint requests to D. J. Polzella, Ph.D., Dept. of Psychology, University of Dayton, 300 College Park, Dayton, OH 45469.

tions of ACM performance in the form of pilot ratings and rankings. The most extensive data were obtained through the automatic recording (at a 2-Hz sampling frequency) of 67 variables from each aircraft. These included aircraft position, altitude, airspeed, orientation, angle of attack, G-forces, fuel flow, control positions and forces applied, various switch settings, and radar/weapons functions.

Kelly, *et al.* applied several univariate and multivariate statistical procedures in an attempt to find the smallest comprehensive set of measures that would discriminate the differences between expert and novice pilots in 1v1 free engagement ACM. It was found, for example, that expert pilots managed energy more efficiently than novice pilots. Thus, expert pilots used the idle position and speedbrake more often and the afterburner less often. In addition, expert pilots made greater use of the roll axis of the aircraft (an indication of greater maneuverability) and generally spent a greater proportion of engagement time in an offensive attitude. Not surprisingly, expert pilots were more likely than novice pilots to achieve a "kill."

A discriminant analysis identified 13 measures that, when properly weighted, could be added together to form a single metric of ACM performance. These measures are listed in Table I. The metric accounted for 51% of the variability in the free engagement performance data and predicted membership in high or low skill groups with 92% accuracy. Despite its predictive power, however, such metrics are inherently limited. A single metric cannot adequately describe the air-to-air combat environment, which is characterized by complex and dynamic relationships among numerous variables. MDS is well-suited for providing just such a description.

METHODS

In order to apply MDS to Kelly, *et al.*'s (5) data, we first needed a matrix of proximities among the various ACM performance measures. Under normal circumstances proximity measures would be obtained directly, e.g., through judgments of relatedness. Since this was not possible, we used as input the matrix of intercorrelations among the 13 measures that formed the basis of their performance metric. This matrix was included (for both expert and novice pilots) as part of Appendix E of their report.

TABLE I. SIMULATED AIR COMBAT MANEUVERING PERFORMANCE MEASURES.

Performance Measure	Definition
Altitude Rate	Mean Absolute Vertical Speed
Roll Rate	Mean Absolute Roll Rate in $° \cdot s^{-1}$
Idle	Number of Times Throttle in Idle Position
Mil Power	Number of Times Throttle in High Mil Position
Afterburner	Number of Times Throttle in Afterburner Position
Speed Brake	Mean Speed Brake Deflection
Fuel Flow	Mean Fuel Flow in $lb \cdot h^{-1}$
Energy Mgt	RMS Energy Management Index
Start Position	Initial Simulator Configuration
Out of View	% Time Out of Opponent's View
Offense	% Time Opponent Positioned at a Sight Angle of Less Than 60°
Gun Range	% Time Opponent Positioned Within Gun Range
Gun Kill	Probability of Gun Kill

The SPSS-X (13) procedure ALSCAL was used to compute separate MDS analyses for high- and low-skilled pilots. ALSCAL uses the alternating least-squares approach to MDS proposed by Young, Takane, and Lewyckyj (16). An extremely flexible procedure, ALSCAL is the only scaling program that incorporates metric, nonmetric, classical, replicated, and weighted (i.e., individual differences) MDS models. For the present application a classical, nonmetric, Euclidean model was invoked. The data were considered to be ordinal (the default) and continuous. Although ALSCAL can theoretically output a configuration of up to six dimensions, the most appropriate dimensionality is constrained by several factors, including the number of stimuli. In this case, with only 13 stimuli, a two-dimensional solution seemed most appropriate.

Interpretation of an MDS configuration can be achieved by several means. The most common approach is to determine if some objective stimulus parameter(s) can be projected through the space. A technique that is frequently used is to regress a set of candidate stimulus parameters (obtained from rating scales) on the set of stimulus coordinates (7). A particular scale is said to provide a satisfactory interpretation of a dimension if the multiple correlation for the scale and the regression weight on that dimension are both high. Unfortunately, the data necessary for performing these analyses were not provided in the Kelly, *et al.* (5) report. An alternative approach is simply to look at the properties or attributes of the stimuli at each end of a dimension and determine whether some property changes in a systematic fashion. (A rotation of the coordinate system is sometimes necessary.)

RESULTS

The derived stimulus configurations for expert and novice pilots are shown in Fig. 1 and 2, respectively. In both cases the correspondence between the relative distances in the configuration and the original proximity measures was extremely high. For the experts' data the configuration accounted for 97.8% of the variability in the proximity matrix. STRESS, a badness-of-fit measure attributed to Kruskal (6), was only 0.060. Similarly, for the novices' data the configuration accounted for 97.0% of the variability in the proximity matrix, with STRESS equal to 0.077.

Expert pilots: In the case of the expert pilots' configuration (Fig. 1), it appears that the horizontal axis represents an energy management (left)—maneuverability (right) dimension. The cluster of stimuli on the left are all energy-related, which indicates that the expert pilots' performance was characterized by frequent throttle activity. The cluster of stimuli on the right are all related to ACM, which indicates that mission success

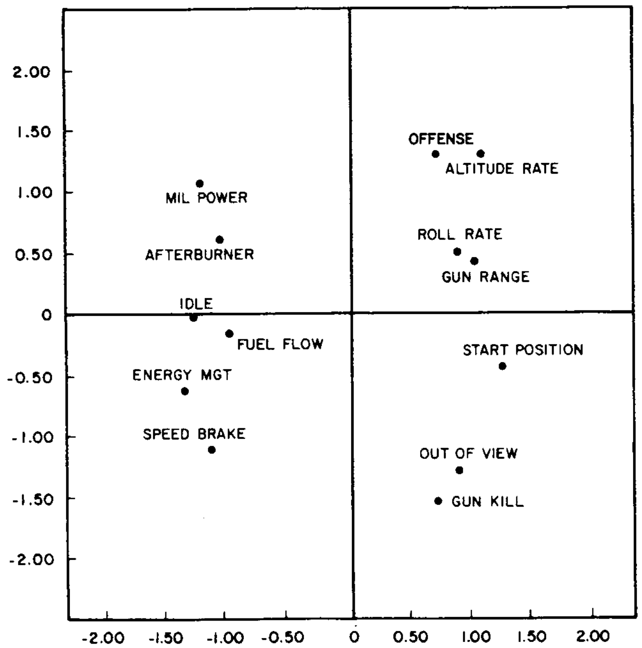

Fig. 1. Derived stimulus configuration for expert pilots.

was associated primarily with offensive and defensive maneuverability. The interpretation of Dimension 2 is not immediately obvious, but careful inspection suggests that it relates to "mission status." Thus, stimuli on the bottom (e.g., gun kill, out-of-opponent's-view, efficient energy management) are direct indicants of mission success, whereas stimuli on the top (e.g., high throttle settings, offensive maneuverability) relate to mission dynamics. [It is important to note that these interpretations are subjective and somewhat arbitrary. The most common technique for "objectively" labelling dimensions was precluded in this situation due to the fact that rating profiles of the ACM performance measures were not a part of the original database. Nevertheless, according to Schiffman, *et al.* (11), despite the fact that no dimensions as such may be found, stimuli may still be arranged in a sensible manner, as was the case here.]

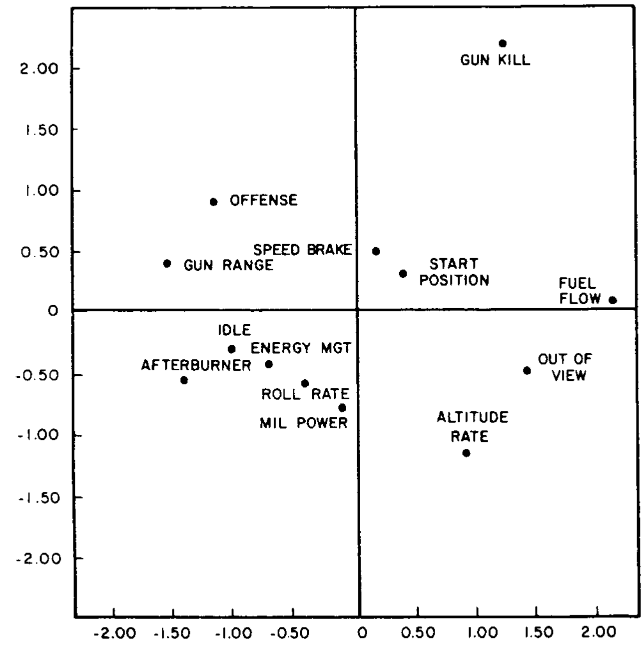

Fig. 2. Derived stimulus configuration for novice pilots.

Novice pilots: In the case of the novice pilots' configuration (Fig. 2), neither dimension is readily interpretable. However, there are marked contrasts between this configuration and that of the experts. The isolated location of gun kill indicates that mission success was not related either to efficient energy management or to skillful ACM performance. In addition, the relative proximity of fuel flow, out-of-view, and altitude rate indicates that a state of defensive advantage was achieved at the expense of increased fuel use. Finally, the presence of offense, gun range, and roll rate within a cluster of energy management stimuli reveals that offensive ACM was more dependent on throttle activity for novice pilots than it was for expert pilots.

DISCUSSION

The results of this study indicate that there are important differences in ACM performance between expert and novice pilots. It is not merely that expert pilots are more skillful. This is to be expected, and any valid performance measure would surely reveal this fact. The present study revealed, in addition, that expert pilots characteristically show advantageous maneuverability and intelligent energy management. What emerges from these results is a view of the air-to-air combat environment that is much more comprehensive than could be provided by a single performance metric. MDS is a powerful analytic tool, which can help the research scientist to discover the complex dynamics of ACM performance.

REFERENCES

1. Beatty J. Mapping the functional proximity of cortical regions by multidimensional scaling. Neurosci. Lett. 1987; 8:99–104.
2. Breidenbach ST, Ciavarelli AP, Sievers R, Lilienthal M. Measurement methods and metrics for aircrew assessment during close-in air-to-air combat. Orlando, FL: Naval Training Systems Center, Naval Air Systems Command. 1985; (NAVAIRSYSCOM N00019-81-C-0098).
3. Clark WC, Carroll JD, Yang JC, Janal MN. Multidimensional scaling reveals two dimensions of thermal pain. J. Exp. Psychol. [Hum. Percept.]. 1986; 12:103–7.
4. Hasenfus N, Martindale C, Birnbaum D. Psychological reality of cross-media artistic styles. J. Exp. Psychol. [Hum. Percept.]. 1983; 9:841–63.
5. Kelly MJ, Wooldridge L, Hennessy RT, Vreuls D, Barnebey SF, Cotton JC, Reed JC. Air combat maneuvering performance measurement. Williams Air Force Base, AZ: Flying Training Division, Air Force Human Resources Laboratory. 1979; (NAVTRAEQUIPCEN IH 315/AFHRL-TR-79-3).
6. Kruskal JB. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrika. 1964; 29:1–27.
7. Kruskal JB, Wish M. Multidimensional scaling. Beverly Hills, CA: Sage Publications, 1978.
8. Laxar K, Moeller G, Rogers WH. The cognitive organization of submarine sonar information: a multidimensional scaling analysis. Groton, CT: Naval Submarine Medical Research Laboratory. 1983; (NSMRL Report No. 1010).
9. Medlin SM, Thompson P. Evaluator rating of unit performance in field exercises: a multidimensional scaling analysis. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. 1980; (Technical Report 438).

10. Sands SF, Lincoln CE, Wright AA. Pictorial similarity judgments and the organization of visual memory in the rhesus monkey. J. Exp. Psychol. [Gen.]. 1982; 111:369–89.
11. Schiffman SS, Reynolds ML, Young FW. Introduction to multidimensional scaling: theory, methods, and applications. New York: Academic Press. 1981.
12. Schvaneveldt RW, Goldsmith TE, Durso FT, Maxwell K, Acosta HM, Tucker RG. Structures of memory for critical flight information. Williams Air Force Base, AZ: Operations Training Division, Air Force Human Resources Laboratory, 1982; (AFHRL-TP-81-46).
13. SPSS-X user's guide (2nd ed.). New York, NY: McGraw-Hill 1986.
14. Toye RC. The effect of viewing position on the perceived layou of space. Percept. Psychophys. 1986; 40:85–92.
15. Wish M. Notes on the variety, appropriateness, and choice o *proximity measures. Paper prepared for the workshop on mul*tidiminsional scaling. Murray Hill, NJ: Bell Laboratories 1972; June 7–10.
16. Young FW, Takane Y, Lewyckyj R. ALSCAL: a nonmetric multidimensional scaling program with several differences options Behav. Res. Meth. Instrum. 1978; 10:451–3.